POLSKA AKADEMIA NAUK

Instytut Badań Systemowych

WSPOMAGANIE INFORMATYCZNE ROZWOJU SPOŁECZNO-GOSPODARCZEGO I OCHRONY ŚRODOWISKA

Redakcja:
Jan Studziński
Ludosław Drelichowski
Olgierd Hryniewicz

WSPOMAGANIE INFORMATYCZNE
ROZWOJU
SPOŁECZNO-GOSPODARCZEGO
I OCHRONY ŚRODOWISKA

Polska Akademia Nauk Instytut Badań Systemowych

Seria: BADANIA SYSTEMOWE

tom 36

Redaktor naukowy:

Prof. dr hab. Jakub Gutenbaum

Warszawa 2004

WSPOMAGANIE INFORMATYCZNE ROZWOJU SPOŁECZNO-GOSPODARCZEGO I OCHRONY ŚRODOWISKA

Redakcja:

Jan Studziński
Ludosław Drelichowski
Olgierd Hryniewicz

Książka wydana dzięki dotacji KOMITETU BADAŃ NAUKOWYCH

Książka zawiera wybór artykułów poświęconych omówieniu aktualnego stanu badań w kraju w zakresie rozwoju modeli, technik i systemów zarządzania oraz ich zastosowań w różnych dziedzinach gospodarki narodowej. Wyodrębnioną grupę stanowią artykuły omawiające aplikacyjne wyniki projektów badawczych i celowych KBN.

Recenzenci artykułów:

Dr Lucyna Bogdan
Prof. dr hab. inż. Olgierd Hryniewicz
Dr Grażyna Petriczek
Prof. dr hab. inż. Andrzej Straszak
Dr inż. Jan Studziński

Komputerowa edycja tekstu: Anna Gostyńska

Wydawca: Instytut Badań Systemowych PAN
ul. Newelska 6, 01-447 Warszawa

Sekcja Informacji Naukowej i Wydawnictw IBS PAN
tel. 836-68-22

Druk: Zakład Poligraficzny Urzędu Statystycznego w Bydgoszczy
Nakład 110 egz.

ISBN 83-85847-92-8
ISSN 0208-8028

POMIAR SIŁY ASOCJACJI SKATEGORYZOWANYCH CECH STATYSTYCZNYCH W PRZYPADKU NIEPRECYZYJNIE ZDEFINIOWANYCH KATEGORII

Olgierd Hryniewicz
Instytut Badań Systemowych, Polska Akademia Nauk
<hryniewi@ibspan.waw.pl>

Measures of the strength of association (dependence) of categorical data are needed in many applications related to the data analysis such as, for example, knowledge discovery. Well known statistical tests of independence, such as Pearson's chi-square test, are not useful for this purpose. However, statistics proposed by Goodman and Kruskal for ordered categorical data may be applied in such a case. In the paper we propose the extension of the Goodman-Kruskal gamma statistic for the case of multiple answers that are present in the data due to imprecise definitions of categories.

Keywords: Categorical data, measures of association, Goodman-Kruskal gamma statistic, multiple answers, imprecise categories.

1. Wprowadzenie

Badanie zależności pomiędzy cechami statystycznymi jest jednym z najważniejszych zadań statystyki. Rozpatrywane są zazwyczaj dwie grupy zagadnień. Pierwsza z nich obejmuje metody analizy zależności występujących pomiędzy cechami statystycznymi ciągłymi. Znajdują tu zastosowanie wszelkiego rodzaju analizy korelacyjne. Do drugiej grupy zaliczamy metody analizy zależności pomiędzy cechami statystycznymi dyskretnymi. Szczególnie interesujący jest tu przypadek, gdy analizowane cechy przyjmują wartości należące do pewnego skończonego (zwykle niezbyt licznego) zbioru kategorii. Mówimy w takim przypadku o analizie zależności pomiędzy skategoryzowanymi cechami statystycznymi. Przypadek ten obejmuje również sytuacje analizy zależności zmiennych jakościowych. Zadanie to jest szczególnie istotne w obszarze nauk o zarządzaniu oraz nauk ekonomicznych, a także wszędzie tam, gdzie nie jest możliwy dokładny pomiar analizowanych cech statystycznych. Ze zjawiskiem tym mamy zwłaszcza do czynienia, gdy źródłem wyniku pomiaru jest człowiek. W takim przypadku wynik pomiaru cechy statystycznej sprowadza się tyko do wskazania odpowiedniej kategorii, co jest rzeczą znacznie prostszą od określenia tego wyniku np. na skali liczbowej.

W rachunku prawdopodobieństwa pojęcie zależności nie jest zdefiniowane w sposób jednoznaczny. Potrafimy natomiast jednoznacznie zdefiniować pojęcie niezależności. W związku z tym, w literaturze statystycznej znacznie więcej miejsca poświęca się problemowi weryfikacji hipotez o niezależności badanych cech statystycznych, niż problemowi pomiaru siły zależności występujących pomiędzy analizowanymi cechami statystycznymi. Brak jednej definicji zależności skłania również statystyków do mówienia raczej o pomiarze siły *asocjacji* (*związku*) pomiędzy analizowanymi cechami, niż o pomiarze stopnia (siły) zależności. Z punktu widzenia praktyka rozróżnienie to nie ma specjalnego znaczenia, a pojęcia pomiaru siły asocjacji oraz siły zależności mogą być traktowane jako zamienne.

W przypadku skategoryzowanych cech statystycznych wyniki obserwacji (badań) przedstawiane są w postaci tzw. tablicy korelacyjnej. Rozpatrzmy przypadek analizy zależności dwu cech jakościowych, których wartości dzielimy na kilka *kategorii*. Załóżmy, że cecha jakościowa X przyjmuje wartości należących do k rozłącznych kategorii oznaczonych przez x_1, x_2,...,x_k, a cecha jakościowa Y przyjmuje wartości należące do r rozłącznych kategorii oznaczonych przez y_1, y_2,...,y_r. Dane statystyczne mają wówczas postać dyskretną i można je przedstawić w postaci następującej tablicy

Tablica 1. Tablica korelacyjna dwu cech jakościowych

x_i / y_j	y_1	y_2	...	y_j	....	y_r	Suma
x_1	n_{11}	n_{12}	...	n_{1j}	...	n_{1r}	$n_{1.}$
x_2	n_{21}	n_{22}	...	n_{2j}	...	n_{2r}	$n_{2.}$
....	....	....	...	....	....	....	....
x_i	n_{i1}	n_{i2}	...	n_{ij}	...	n_{ir}	$n_{i.}$
...	...	...	...	...	...	...	...
x_k	n_{k1}	n_{k2}	...	n_{kj}	...	n_{kr}	$n_{k.}$
Suma	$n_{.1}$	$n_{.2}$	...	$n_{.j}$	...	$n_{.r}$	$n_{..}$

W tablicy korelacyjnej (zwanej też *tablicą dwudzielczą* lub *tablicą kontyngencji*) wielkość n_{ij} oznacza liczbę obserwacji, dla których wartości cechy jakościowej X należącej do i-tej kategorii odpowiada wartość cechy jakościowej Y należąca do j-tej kategorii.

Dla danych statystycznych przedstawionych w postaci tablicy korelacyjnej opracowano szereg testów statystycznych służących do weryfikacji hipotezy o *statystycznej niezależności* analizowanych cech. Najczęściej stosowany w praktyce test chi-kwadrat Persona jest opisany praktycznie we wszystkich podręcznikach statystyki. Statystyka testowa ma w przypadku tego testu postać

$$\chi^2 = \sum_{i=1}^{k}\sum_{j=1}^{r}\frac{\left(n_{ij}-\hat{n}_{ij}\right)^2}{\hat{n}_{ij}} = \sum_{i=1}^{k}\sum_{j=1}^{r}\left(\frac{n_{ij}^2}{\hat{n}_{ij}}\right) - n \tag{1}$$

gdzie

$$\hat{n}_{ij} = \frac{n_{i.}n_{.j}}{n}, i = 1,\dots k, j = 1,\dots r \tag{2}$$

$$n_{i.} = \sum_{j=1}^{r} n_{ij}, i = 1,\dots,k \tag{3}$$

$$n_{.j} = \sum_{i=1}^{k} n_{ij}, j = 1,\dots,r\,, \tag{4}$$

Na bazie statystyki chi-kwadrat (1) skonstruowano wskaźniki, które mogą być wykorzystywane do pomiaru siły zależności. Najbardziej znanym z nich jest współczynnik zbieżności Czuprowa określony wzorem

$$T = \sqrt{T_{xy}^2} \tag{5}$$

gdzie

$$T_{xy}^2 = T_{yx}^2 = \frac{\chi^2}{n\sqrt{(r-1)(k-1)}}\,. \tag{6}$$

Jeżeli wartość statystyki *T jest bliska zeru,* to możemy mówić o *niezależności* badanych zmiennych losowych, zaś w przypadku gdy zachodzi równość T=1 mamy przypadek zależności *funkcyjnej*. W statystyce używane jest również pojęcie współczynnika determinacji Czuprowa równego $100 * T_{xy}^2$, który mówi w ilu procentach zmienność zmiennej zależnej Y jest określona zmiennością zmiennej niezależnej X. Innym wskaźnikiem siły asocjacji jest współczynnik V Cramera zdefiniowany jako

$$V = \sqrt{\frac{\chi^2}{n(m-1)}} \tag{7}$$

gdzie

$$m = \min(k, r) \tag{8}$$

Współczynniki te, a także inne im podobne, przyjmują wartości z przedziału [0,1] i ich większe wartości interpretowane są jako wskaźniki silniejszej zależności (związku) występującej pomiędzy analizowanymi cechami statystycznymi. Niestety, maja one pewną istotną wadę. Jest nią brak interpretacji probabilistycznej. Oznacza to, że dany współczynnik nie ma interpretacji prawdopodobieństwa zajścia jakiegoś zdarzenia. Brak takiej interpretacji uniemożliwia wnioskowanie statystyczne o sile interesującego nas związku pomiędzy badanymi cechami.

Wspomnianej powyżej wady nie mają miary asocjacji wprowadzone przez Goodmana i Kruskala (1954), a także inne, podobnie zdefiniowane, miary siły asocjacji (związku). Są one powszechnie stosowane w praktyce statystycznej (ze względu na dostępność w powszechnie stosowanych komputerowych pakietach statystycznych takich jak SPSS lub STATISTICA), natomiast bardzo rzadko opisywane w polskiej literaturze statystycznej. Do chlubnych wyjątków należy tu książka Koronackiego i Mielniczuka (2001). Wykorzystaniu jednego z tych mierników – statystyki γ Goodmana-Kruskala poświęcona jest niniejsza praca. W rozdziale drugim przypomnimy definicje tej statystyki oraz jej podstawowe własności. Z kolei, w rozdziale trzecim przedstawimy nowe wyniki dotyczące wykorzystania statystyki γ Goodmana-Kruskala w przypadku nieprecyzyjnie określonych kategorii. Rozpatrzymy przypadek, gdy brak precyzji w definiowaniu poszczególnych kategorii prowadzi do niejednoznacznych odpowiedzi. Ograniczymy się wyłącznie do przypadku odpowiedzi wielokrotnych. Pracę zakończy krótkie podsumowanie i omówienie możliwych uogólnień przedstawionego problemu.

2. Pomiar siły asocjacji z wykorzystaniem statystyki γ Goodmana-Kruskala

Propozycje różnych mierników siły asocjacji skategoryzowanych cech statystycznych pojawiały się już w czasach, gdy statystyka nie była wydzieloną gałęzią nauki. Z próbami rozwiązania tego problemu związane są najważniejsze nazwiska w historii statystyki. Przełomowym momentem stało się jednak opublikowanie pracy Goodmana i Kruskala (1954), którzy zaproponowali jednolitą metodologię podejścia do tego problemu. Praca ta, a także inne prace tej dwójki autorów legły u podstaw serii propozycji nowych mierników siły asocjacji, które zostały zaproponowane przez wielu autorów. W polskiej literaturze statystycznej podstawowe informacje na temat tych i podobnych procedur statystycznych można znaleźć w książce Koronackiego i Mielniczuka (2001). W niniejszej pracy omówimy bliżej tylko jedną z zaproponowanych przez Goodmana i Kruskala procedur, a mianowicie pomiar siły asocjacji wykorzystujący statystykę γ Goodmana-Kruskala dla przypadku, gdy kategorie analizowanych cech statystycznych są uporządkowane.

W ogólnym przypadku analizy danych dotyczących cech statystycznych o wartościach skategoryzowanych (np. cech jakościowych) poszczególne kategorie należy traktować jako *etykiety*. Tak też są one traktowane we wspomnianym już teście niezależności chi-kwadrat Pearsona. Przyjęcie takiego założenia oznacza, że wynik analizy nie zależy od uporządkowania tych wartości w tablicy korelacyjnej. W wielu jednak przypadkach występuje naturalne uporządkowanie kategorii rozpatrywanych cech. Ma to, na przykład, miejsce w przypadku, gdy kategorie odpowiadają kolejnym przedziałom wartości analizowanej cechy. Innym przykładem są kategorie odpowiadające różnym przypadkom „intensywności" występowania analizowanej cechy. Na przykład, w ankietowych badaniach marketingowych dotyczących oceny skłonności do zakupu określonego produktu lub usługi możemy oczekiwać od osób ankietowanych wyboru jednej z trzech opcji: „nie kupię", „być może kupię" oraz „zdecydowanie kupię". Z kolei, osoby ankietowane mogą określać siebie jako należące do grupy osób o dochodach należących do określonego przedziału liczbowego. Informacja o uporządkowaniu kategorii rozpatrywanych cech statystycznych pomaga nam określić siłę asocjacji (związku) pomiędzy analizowanymi cechami.

Zanim wprowadzimy statystykę γ Goodmana-Kruskala przyjmijmy dwie definicje.

Definicja 1 (Koronacki, Mielniczuk, 2001)

Mówimy, że para jednostek jest ***zgodna****, jeżeli jest spełniony jeden z następujących warunków:*

(*1*) *ranga kategorii, jaką przyjmuje zmienna X w przypadku pierwszej jednostki w parze, jest wyższa od rangi kategorii, jaką przyjmuje zmienna X dla drugiej jednostki w parze, oraz ranga kategorii, jaką przyjmuje zmienna Y dla pierwszej jednostki w parze, jest wyższa od rangi kategorii, jaką przyjmuje zmienna Y dla drugiej jednostki w parze jednostek.*

(2) *ranga kategorii, jaką przyjmuje zmienna X w przypadku pierwszej jednostki w parze, jest niższa od rangi kategorii, jaką przyjmuje zmienna X dla drugiej jednostki w parze, oraz ranga kategorii, jaką przyjmuje zmienna Y dla pierwszej jednostki w parze, jest niższa od rangi kategorii, jaką przyjmuje zmienna Y dla drugiej jednostki w parze jednostek.*

Definicja 2 (Koronacki, Mielniczuk, 2001)

Para jednostek jest ***niezgodna****, jeżeli jedna z jednostek w parze ma kategorię zmiennej X wyższej rangi oraz kategorię zmiennej Y niższej rangi niż druga z jednostek pary.*

Goodman i Kruskal (1954) jako miarę siły asocjacji zaproponowali odpowiednio znormalizowaną różnicę prawdopodobieństw, że losowo wybrana para jednostek jest zgodna i niezgodna. Oszacowaniem tego wskaźnika jest statystyka γ Goodmana-Kruskala zdefiniowana przy pomocy wzoru

$$\gamma = \frac{C - D}{C + D} \tag{9}$$

gdzie C jest liczbą wszystkich par zgodnych w tablicy korelacyjnej, zaś D jest liczbą wszystkich par niezgodnych. Wielkości te można w sposób formalny zapisać w postaci

$$C = \sum_{i,j} n_{ij} C_{I;ij} \tag{10}$$

gdzie

$$C_{I;ij} = \sum_{i'>i} \sum_{j'>j} n_{i'j'} \tag{11}$$

oraz

$$D = \sum_{i,j} n_{ij} D_{IV;ij} \tag{12}$$

gdzie

$$D_{IV;ij} = \sum_{i'>i} \sum_{j'<j} n_{i'j'} \tag{13}$$

Wartości statystyki γ Goodmana-Kruskala bliskie zera świadczą o słabej zależności pomiędzy analizowanymi cechami (w idealnym przypadku niezależności wartość statystyki jest równa zeru). Z kolei, wartości bliskie jeden (minus jeden) świadczą o bardzo silnym dodatnim (ujemnym) związku pomiędzy analizowanymi cechami statystycznymi. Goodman i Kruskal (1972) analizowali również asymptotyczny rozkład prawdopodobieństwa statystyki γ. Znajomość tego rozkładu pozwala wnioskować o istotności statystycznej zaobserwowanej miary siły asocjacji.

Sposób wyznaczania wartości statystyki γ Goodmana-Kruskala zilustrujmy na przykładzie hipotetycznej analizy zależności intensywności palenia papierosów od wykształcenia osób ankietowanych. Przyjmijmy, że wyniki badania przedstawione są w następującej tablicy korelacyjnej

Tablica 2. Tablica korelacyjna przedstawiająca wyniki badań

	Nie pali	Pali umiarkowanie	Pali dużo	Razem
Podstawowe	15	10	10	35
Średnie	25	10	5	30
Wyższe	30	10	10	50
Razem	70	30	25	125

Dla powyższych danych mamy:

$$C = 15*(10+5+10+10)+10*(5+10)+25*(10+10)+10*10 = 1275$$

$$D = 10*(25+30)+10*(25+10+30+10)+10*30+5*(30+10) = 1800$$

i wobec tego

$$\gamma = \frac{1275-1800}{1275+1800} = -0{,}171$$

Wyniki powyższego badania świadczą o występowaniu niezbyt mocnej ujemnej zależności pomiędzy wykształceniem a intensywnością palenia papierosów (im wyższe wykształcenie, tym badane osoby paliły w sposób mniej intensywny lub też częściej nie paliły w ogóle).

3. Pomiar siły zależności w przypadku niejednoznacznych informacji z badań

W rozpatrywanym powyżej przypadku przyjmuje się podstawowe założenie, że poszczególne obserwacje są określone w sposób jednoznaczny przez podanie kategorii odpowiadającej zmiennej X oraz kategorii odpowiadającej zmiennej Y. Założenie to może być spełnione, gdy poszczególne kategorie zdefiniowane są w sposób jednoznaczny. Jednakże w wielu przypadkach, patrz przykład w punkcie drugim, poszczególne kategorie mogą być określone w sposób nieprecyzyjny. W takich przypadkach może dojść do sytuacji, gdy nie wiadomo, do której kategorii należy zaliczyć daną obserwację. W ogólnym przypadku można przyjąć, że każda obserwacja może być zaliczona do kilku kategorii z określoną miarą przynależności. Na przykład, w pracy Hryniewicz (2004) założono, że kategoria zmiennej X określona jest w sposób jednoznaczny, natomiast kategoria zmiennej Y określona jest w sposób rozmyty.

Każda obserwacja opisana jest parą $\left(X_q, \widetilde{Y}_q\right)$, gdzie X_q jest zaobserwowaną kategorią zmiennej X, zaś $\widetilde{Y}_q = y_1 \mid \mu_{i_q 1} + y_2 \mid \mu_{i_q 2} + \ldots + y_r \mid \mu_{i_q r}$, $q = 1,\ldots,n$, przy czym $\mu_{i_q j} \in [0,1]$, $j = 1,\ldots,r$ jest zaobserwowaną rozmytą kategorią zmiennej Y. W niniejszej pracy, ze względu na jej ograniczony rozmiar, rozpatrzymy prostszy przypadek, gdy $\mu_{i_q j} \in \{0,1\}$, $j = 1,\ldots,r$, tzn. gdy dla danej obserwacji mogą być jednocześnie wskazanych *kilka* kategorii zmiennej Y. W takim przypadku wynik obserwacji zmiennej Y opisany jest wektorem $Y_q^1 = \left\{M_{i_q 1}^1, M_{i_q 2}^1, \ldots, M_{i_q r}^1\right\}$, gdzie

$$M_{i_q j}^1 = \begin{cases} 1 & \text{gdy } \mu_{i_q j} = 1 \\ 0 & \text{w przeciwnym przypadku} \end{cases}, \quad j = 1,\ldots,r, \quad q = 1,\ldots,n.$$

Niech $Y^1 = \{Y_1^1, Y_2^2, \ldots, Y_n^1\}$ oznacza zbiór wszystkich możliwych obserwacji zaś $S^1 \subseteq Y^1$ zbiór takich obserwacji, dla których jedynka występuje tylko dla jednej kategorii. Oznaczmy przez $n_{ij,\min}$ liczbę obserwacji w komórce (i,j) obliczoną po wszystkich elementach zbioru S^1, zaś $n_{ij,\max}$ liczbę obserwacji w komórce (i,j) obliczoną po wszystkich elementach zbioru Y^1. Zdefiniujmy teraz następujący zbiór

$$N^1 = \left\{ n_{ij}, i = 1,\ldots,k,\ j = 1,\ldots,r : n^1_{ij,\min} \le n_{ij} \le n^1_{ij,\max}, \sum_{j=1}^{r} n_{ij} = n_{i\cdot} \right\}.$$

Możemy teraz wyznaczyć przedział możliwych wartości statystyki γ Goodmana-Kruskala z następujących wzorów:

$$\gamma_{\min} = \inf_{N^1} \gamma \tag{14}$$

oraz

$$\gamma_{\max} = \sup_{N^1} \gamma\,. \tag{15}$$

Powyższy wynik należy zinterpretować w następujący sposób. Wskutek nieprecyzyjnych wskazań kategorii zmiennej Y nie jesteśmy w stanie określić dokładnej wartości statystyki γ. Jedyne, co możemy podać to dolną i górną granicę przedziału jej możliwych wartości.

Powyższe rozważania zilustrujmy przykładem, będącym modyfikacją przykładu omówionego w poprzednim punkcie. Przyjmijmy, że dane statystyczne przedstawiają się w sposób następujący:

- 40 osób z wykształceniem średnim i podstawowym określiło się jako osoby niepalące;
- 15 osób z wykształceniem średnim i podstawowym określiło się jako osoby palące umiarkowanie;
- 10 osób z wykształceniem średnim i podstawowym określiło się jako osoby palące dużo;
- 30 osób z wykształceniem wyższym określiło się jako osoby niepalące;
- 8 osób z wykształceniem wyższym określiło się jako osoby palące umiarkowanie;
- 8 osób z wykształceniem wyższym określiło się jako osoby palące dużo;
- 10 osób z wykształceniem średnim i podstawowym określiło się jednocześnie jako osoby palące umiarkowanie oraz osoby palące dużo;
- 4 osoby z wykształceniem wyższym określiło się jednocześnie jako osoby palące umiarkowanie oraz osoby palące dużo.

Dla tak zdefiniowanych danych maksymalna wartość statystyki γ Goodmana-Kruskala uzyskujemy dla następującej tablicy korelacyjnej

Tablica 3. Tablica korelacyjna (maksimum statystyki γ)

	Nie pali	Pali umiarkowanie	Pali dużo	Razem
Średnie i podst.	40	25	10	75
Wyższe	30	8	12	50
Razem	70	33	22	125

Z kolei, minimalna wartość statystyki γ Goodmana-Kruskala odpowiada następującej konfiguracji danych:

Tablica 4. Tablica korelacyjna (minimum statystyki γ)

	Nie pali	Pali umiarkowanie	Pali dużo	Razem
Średnie i podst.	40	15	20	75
Wyższe	30	12	8	50
Razem	70	27	28	125

Odpowiadający powyższym nieprecyzyjnym danym przedział możliwych wartości statystyki γ Goodmana-Kruskala wynosi [-0,1674 , -0,01345]. Jak łatwo zauważyć, brak precyzji w definiowaniu kategorii skutkuje nieprecyzyjnie wyznaczoną wartością statystyki γ Goodmana-Kruskala. Jej możliwe wartości znajdują się blisko zera, co może być interpretowane jako przesłankę by sądzić, że intensywność palenia nie zależy od wykształcenia.

Literatura

Goodman L.A., Kruskal W.H. (1954) Measures of Association for Cross Classification. *Journ. of the Amer. Stat. Assoc.*, **49**, 732–764.

Goodman L.A., Kruskal W.H. (1972) Measures of Association for Cross Classification. IV: Simplification of Asymptotic Variances. *Journ. of the Amer. Stat. Assoc.*, **67**, 414–421.

Hryniewicz O. (2004) Selection of variables for systems analysis – application of a fuzzy statistical test for independence. *Proc. of IPMU'2004*, Perugia.

Koronacki J., Mielniczuk J. (2001) *Statystyka dla studentów kierunków technicznych i przyrodniczych*. WNT, Warszawa.

ISSN 0208-8028
ISBN 83-85847-92-8

W celu uzyskania bliższych informacji i zakupu dodatkowych egzemplarzy
prosimy o kontakt z Instytutem Badań Systemowych PAN
ul. Newelska 6, 01-447 Warszawa
tel. 837-35-78 w. 241 e-mail: biblioteka@ibspan.waw.pl